DRG-Net: Diabetic Retinopathy Grading Network Using Graph Learning with Extreme Gradient Boosting Classifier

*Venkata Kotam Raju Poranki, B. Srinivasa Rao
[1]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India
E-mail: vkrporanki@gmail.com
*Corresponding author

Keywords: color eye fundus images, Diabetic retinopathy, graph convolutional neural network, synthetic minority over-sampling technique, extreme gradient boosting

Received: August 6, 2023

Diabetic retinopathy (DR) is a leading cause of blindness that occurs in different age groups. So, the early detection of DR can save millions of people from blindness issues. Further, the manual analysis of DR requires much processing time and experienced doctors. Hence, computer-aided diagnosis (CAD)-based artificial intelligence models have been developed for an early DR prediction. However, the state-of-the-art methodologies failed to extract the deep balanced features, which resulted in poor classification performance. Therefore, this work implements the DR grading network (DRG-Net) using graph learning properties. Initially, the synthetic minority over-sampling technique (SMOTE) is applied to the EyePACS and Messidor dataset to balance the instances of each DR class into uniform levels. Then, a deep graph correlation network (DGCN) is applied to extract the class-specific features by identifying the relationship. Finally, an extreme gradient boosting (XGBoost) classifier is employed to perform the DR classification with the pre-trained balanced features obtained using SMOTE-DGCN. The obtained simulation results performed on the EyePACS dataset and the Messidor dataset disclose that the proposed DRG-Net resulted in higher performance than state-of-the-art DR grading classification approaches, with accuracy, sensitivity, and specificity of 99.01%, 99.01%, and 98.43% for the EyePACS dataset, respectively, and 99.6%, 99.08%, and 100% for the Messidor dataset.

Povzetek: Članek opisuje novo metodo DRG-Net, ki uporablja grafovno učenje in ekstremno gradientn spodbujevalno učenje za zgodnje odkrivanje diabetične retinopatije.

1 Introduction

Humans with DR pose a danger to their eyesight is expected to total 103.12 million besides 28.54 million, respectively, in the year 2021. The numbers are expected to rise to 160.50 million and 44.82 million by 2045 [1]. Additionally, in underdeveloped nations without access to basic healthcare facilities or a scarcity of ophthalmologists. The underserved regions of the developed world are likewise affected by this issue. Thus, early detection and routine screening may cut the chance of vision loss to 57.0% while also lowering treatment costs. The two major processes in DR detection are screening and diagnosis [2]. Fine pathognomonic DR signals are generally established for this purpose following dilation of pupils (mydriasis). For accurate diagnosis, early identification, patient education, and treatment planning, clear pathognomonic signals are essential. This improves healthcare workers' capacity to recognize and treat diabetic retinopathy. Then, DR screening is performed using slit lamp bio-microscopy with a + 90.0 D lens and direct indirect ophthalmoscopy [3]. Finding lesions associated with DR and comparing them to the criteria for the grading system allows one to diagnose DR. Figure 1 depicts the normal retinal vision and DR-related anomalies. Diabetes complications, such as diabetic macular edema (DME) and DR, are more common in those who are working age.

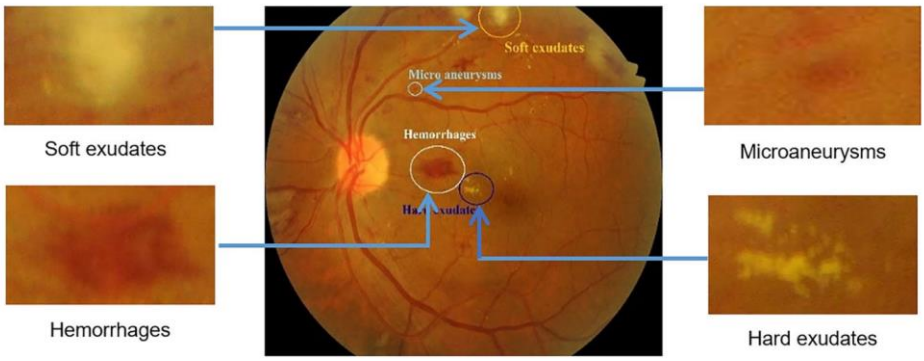


Figure 1: Normal retinal image and abnormalities.

The complicated condition known as DR makes the retinal veins expand and leak blood and fluid [5]. DR causes an issue with visual impairment. Microvascular damage, blood vessel irregularities, inflammation, VEGF, ischemia, compromised blood-fluid barrier, hyperglycemia-induced metabolic alterations, and oxidative stress are some of the variables that contribute to diabetic retinopathy as well as produce retinal vein expansion and leaking. Effective electroretinography transformation, retinal vein type, and retinal blood flow may be used to provide a prompt identification of DR. Clear pathognomonic signals enhance patient care and management by improving diagnostic processes, communication, patient education, severity evaluation, treatment planning, early diagnosis, progression monitoring, and research breakthroughs for diabetic retinopathy. For the early diagnosis of diabetic retinopathy, electroretinography (ERG) is essential because it can spot tiny retinal changes before symptoms manifest. It aids in the distinction of illness phases, forecasts the likelihood of progression, tracks alterations, supports research, and directs individualized therapies. Diabetes patients are more prone to acquire DR if they experience unfavorable symptoms of the condition over an extended length of time [6]. During patients to be examined and during the initial phase of DR therapy to limit the possibility of vision loss, routine retinal screening is essential. DR can be identified by the presence of several signals on a retinal picture. Additionally, these DR symptoms include hard exudates (EX), hemorrhages (HM), soft exudates (EX), and microaneurysms (MA). A diagnostic test used to assess retinal reactions to light, the electroretinogram (ERG) looks for abnormalities that might be signs of diabetic retinopathy (DR). It aids with early discovery, DR severity evaluation, and prompt action. Based on the presence of these symptoms, the DR is divided into five phases: severe DR, no DR, moderate DR, mild DR, and proliferative DR (PDR), which are succinctly described in Table 1 [7]. Non-proliferative DR (NPDR) is used here to refer to all types of DR, including none, mild, moderate, and severe DR. Close examination and monitoring are crucial for patients in the mild stage of DR because they have at least one MA; for patients in the moderate stage who may have HM or MAs in one to three retinal quadrants; for patients in the severe stage who have intraretinal HM and venous beading in two or more quadrants; and for patients in the proliferative stage who need to see a doctor every six to eight months.

Table 1: Various grades of DR are based on signs.

DR severity grading	Class	Signs
No DR	DR-0	Absence of various problems.
Mild DR	DR-1	The presence of MA.
Moderate DR	DR-2	Increase in MA count
Severe DR	DR-3	The intraretinal HM count raises more than twenty. Several quadrants of the vena cava show beading.
Proliferative DR	DR-4	Significant intraretinal microvascular problem with neovascularization in more than one quadrant. Pre-retinal and vascular HM.

The diagnosing stage is currently carried out manually. This technique is expensive, time-consuming, and calls for professionals who have undergone extensive training and have excellent diagnostic skills. Even with all

these tools at hand, a misdiagnosis is still a possibility. The situation is difficult because of this reliance on manual review. Among the often-employed DR grading systems is the Early Treatment Diabetic Retinopathy Study (ETDRS) [10]. The ETDRS isolates highly precise DR characteristics using several layers. This system of grading applies to all seven retinal fundus Fields of View (FOV). Early diagnosis of eye disorders, monitoring health changes, avoiding vision loss, assessing systemic health issues, individualized treatment planning, patient education, and enhanced quality of life are all benefits of routine retinal exams. Even though ETDRS is the gold standard, due to implementation difficulty and technological limitations, alternate systems for grading are also used. One such system is the International Clinical Diabetic Retinopathy (ICDR) [11] scale, which is recognized in both clinical and Computer-Aided Diagnosis (CAD) environments. However, the conventional CAD methods [12] failed to result in maximum performance. So, this work is focused on the development of the following models:

- The SMOTE method is adopted to perform the balancing of instances presented in EyePACS and Messidor datasets, which balances the instances of each DR class into a uniform level.
- The DGCN i.e., graph learning method is developed, and used to extract the class-specific features by identifying the relationship.
- The XGBoost classifier is trained with balanced features and performs the multi-class prediction operation.

The remainder of the text is structured as follows: The overview of DR grading and detection techniques is included in section 2. The planned DRG-Net's analysis is covered in Section 3 along with a description of the sub-modules. The performance comparison with cutting-edge techniques is covered in Section 4 along with a full examination of the simulation results suggested by DRG-Net. The item is concluded in Section 5 with potential future application.

2 Related work

This section gives a detailed analysis of DR grading methods. In [12] authors presented the deep convolutional neural network (DCNN) model for DR classification. Dataset preparation begins with phases for data collecting and data annotation. Furthermore, a median filter was used for data pre-processing. Additionally, DR diagnosis is carried out using DCNN, which categorizes the classes as Normal, Moderate, Heavy, and Severe. This technique is also used in hospitals to provide patients with online services. In [13] authors explored an issue of automatic DR identification and suggested a new deep-learning hybrid to handle. It builds the hybrid model by adding a custom block of CNN layers on top of the pre-trained Inception-ResNet-v2 and using transfer learning on that network.

In [14] authors used single color fundus image and suggested a novel automated deep learning-based technique for severity identification. The suggested method is to build a visual embedding using the DenseNet169 encoder. A convolutional Block Attention Module (CBAM) is also added on top of the encoder to boost its ability to discriminate. To improve network performance and model comprehension, the Convolutional Block Attention Module (CBAM) combines channel and spatial attention processes to boost feature identification in convolutional neural networks. However, this method has high loss values compared to machine learning models. In [15] authors used a fine-tuned ResNet-50, which is developed using two-stage deep learning models. The architecture can effectively classify DR into one of three categories (normal, moderate DR, and severe DR). Further classification of moderate DR was performed using a fine-tuned ResNet-18 [15], while classification of severe DR was performed using a fine-tuned ResNet-50.

In [16] authors proposed the creation and Optimization CNN (OCNN) Model. The pre-trained OCNN model is initially used to extract the features. A classification layer based on gradient boosting is added to enhance features. The examination of the suggested system using a 10-fold cross-validation on two difficult problems shows that it performs better than cutting-edge approaches. In [17] authors proposed DR early detection based on multifractal geometry. Furthermore, automating the diagnostic process and increasing the resulting accuracy by employing a supervised machine learning technique like the Support Vector Machine (SVM) algorithm. In [18] authors proposed a new technique to identify DR using ensemble recurrent neural networks (ERNN). First, a two-stage classifier that uses an assembly approach to merge several machine learning algorithms for categorization. By merging separate predictions and applying a meta-classifier for final judgment, the two-stage classifier assembly strategy, which incorporates numerous machine learning algorithms, improves categorization accuracy. The classifier is also used with DR. So, the issue is that it takes a long time to diagnose this condition, even though an early diagnosis is necessary to prevent total blindness. In [19] authors proposed an

early detection of DR using Principal Component Analysis (PCA) with Firefly optimization. Then, deep learning models were used to perform grading operations. Here, the raw dataset is first normalized using the standard scalar approach, and the most important characteristics in the dataset are then extracted using PCA. Additionally, the Firefly algorithm is used to reduce the number of dimensions. In [20] authors developed the deep belief neural network (DBNN) to predict DR. Here, the Grey Wolf Optimization (GWO) method is used to optimize the PCA features. Additionally, using GWO makes it possible to choose the best training parameters for the DBNN model.

Various machine learning models like support vector machine (SVM) [21], decision tree, and random forest, are used to classify the DR grades from eye fundus images. Eye fundus pictures are entered into a connected graph approach, which represents them as a graph with nodes and edges. With the use of Graph Neural Networks, which analyze the graph and extract characteristics and contextual data, eye ailment diagnoses are now more accurate. Further, the application of different pre-trained convolutional neural network (CNN) [22] architectures was implemented to identify the DR from fundus pictures with rejection resampling (random under-sampling at mini-batch level) technique for effective performance in tackling different imbalanced scenarios of varied dataset sizes. To manage unbalanced datasets, a variety of strategies are utilized, including resampling, algorithm-level, algorithmic, synthetic data creation, transfer learning, and hybrid methods. Experimentation is necessary for the best outcomes. Further, the classification of DR using the Ensemble of Machine Learning (EML) [23] classifier is implemented. However, this method contains higher computational complexity. Further, three-dimensional semantic segmentation of DR lesions with grading operation is implemented using ResNet [24] based transfer learning. The marine predictor algorithm (MPA) is used to select the best features from available features. To address these shortcomings of traditional deep learning models, the explainable and interpretable diabetic retinopathy (ExplainDR) [25] grading approach was developed. Here, neural-symbolic learning is introduced for class-specific feature extraction. Then, a model called the Hinge Attention Network (HA-Net) [26] was created for a thorough study of several DR grades. The HA-Net uses the VGG16 model [27] for feature extraction and grading done by long short-term memory (LSTM). Then, to improve the DR findings, the residual attention network (RAN) [28] was invented. The attention mechanism performs the dilated convolution for feature analysis. In addition, DR diagnosis is carried out using transfer learning based InceptionResNetV2 [29], Xception, and EfficientNetB3 model. Here, InceptionResNetV2 resulted in superior performance. Then, Mask Region-based CNN (MRCNN) [30] was utilized instead of faster region-based CNN (Faster RCNN) and transferred learning model to classify DR grades [31].

3 Proposed system

This section gives a detailed analysis of the proposed DRG-Net. The DRG-Net is a hybrid learning model, that is developed by the properties of graph learning, deep learning, and machine learning models. Figure 2 shows the block diagram of the proposed DRG-Net. However, the EyePACS and Messidor datasets contain the five classes of DR, i.e., DR-0, DR-1, DR-2, DR-3, DR-4. However, the number of images in each grade is uniform, which results in imbalanced features. So, the SMOTE method is applied to balance the number of instances in each class, which maintains an equal number of features. Then, the DR contains the highly correlated features among PDR and NPDR classes. So, DGCN is applied to extract the DR grade-specific features by adopting the graph interconnections of features from SMOTE balanced features. Then, the XGBoost classifier is trained with DGCN balanced features and performs classification of multiple DR grades. Due to its excellent predictive strength and capacity to recognize complicated patterns, the Deep Graph Convolutional Network-trained XGBoost classifier excels at categorizing different diabetic retinopathy grades, although its success is reliant on factors like data variety and practitioner knowledge.

3.1 SMOTE data balancing

The dataset's irregular sizes hurt the performance of deep learning classifiers. The datasets have distinct, imbalanced numbers in them. Therefore, most of the DR class features will be given the highest probabilities (priorities) by the XGBoost classifier, while the minority DR class features will be ignored. Using weighted loss functions, sampling techniques, and adaptable goal functions, XGBoost applies ways to ensure equitable minority class characteristics in the classification of unbalanced datasets. Misclassifications are minimized, and problems are addressed. This may lead to inaccurate categorization and inaccurate forecasting. The block diagram for the SMOTE data balancing approach is shown in Figure 3. It is an oversampling strategy that consistently raises the number of minority-class records. The oversampling procedure in this case generates duplicate samples using the closest possible combinations. The Synthetic Minority Over-Sampling Technique (SMOTE), which ensures variety while keeping patterns and carefully

chooses k nearest neighbors, is used in oversampling to create duplicated samples from minority class instances.

The SMOTE method determines the overall dataset record count as well as the record count for every class. SMOTE is a strategy that enhances model performance by creating synthetic records for minority classes, recognizing them, figuring out the oversampling ratio, and arbitrarily choosing the nearest neighbors. Unbalanced data can impair DRG-Net's effectiveness by introducing biases, lowered sensitivity, false negatives, and overfitting. To create synthetic minority class samples, SMOTE is essential. The average record value will then be calculated. Then, SMOTE uses the notion of randomization to create KNNs for every data. Consider the dataset with X minority class samples. Then, $y_1, y_2, \cdots, and\ y_N$ represent the N showed samples for each minority class. For model performance, SMOTE seeks to provide a balanced dataset; however, normalizing the dataset to equal samples defeats this goal. It is advised to preserve or lessen the class disparity when taking SMOTE. So, there must be more than N-number of KNN samples, i.e., $K > N$. Following this, using the correlation among records, random samples were added to the database. The random correlation factor p_i is constructing the interpolation operation.

$$p_i = X + rand(0,1) \times (y_i - X), \qquad i = 1,2,\cdots,N \tag{1}$$

Here, it is expected that the random integer will fall between [0,1]. The y_i denotes the nearest neighbor for sample i, the X represents data records presented in each minority class. By performing the rounding operation to imbalanced minority class X_{IL}, dataset balance was achieved. Due to biases, restricted minority learning, and other problems, deep learning classifiers have difficulty with unbalanced dataset sizes. Oversampling and under-sampling are rebalancing approaches that enhance the accuracy of classifiers for both classes.

$$X_B = round(X_{IL}) * p_i \tag{2}$$

Here, X_B represents the balanced dataset. Lastly, SMOTE equalizes the whole feature set by inserting these random records into the minority class determined by KNN resemblance.

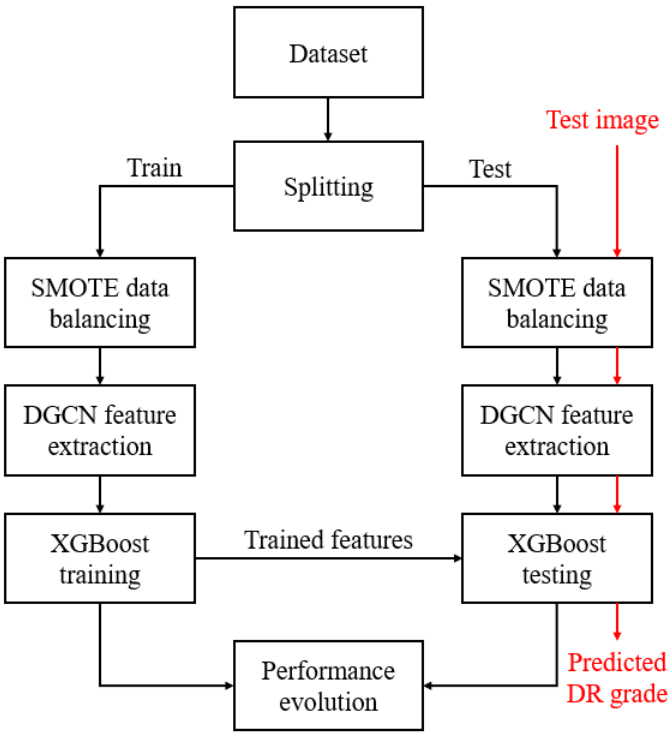


Figure 2: Proposed DRG-Net block diagram.

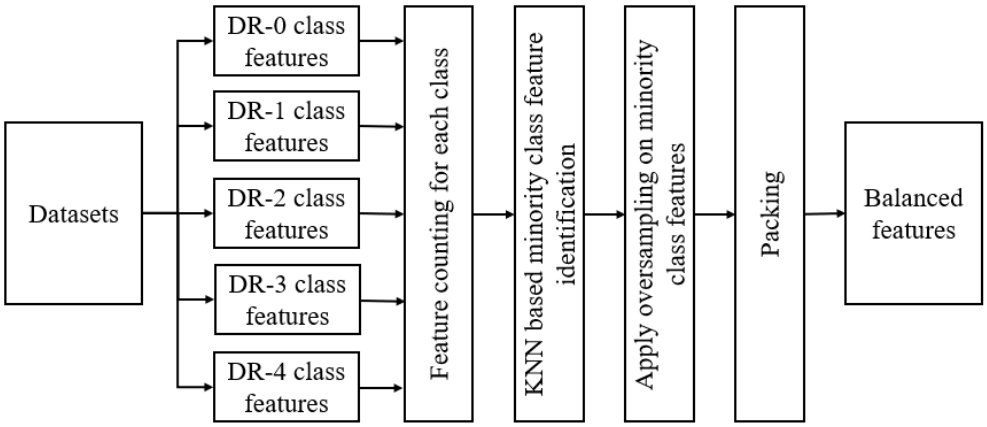


Figure 3: Block diagram of data balancing with SMOTE.

3.2 DGCN-based feature extraction

Graph neural networks (GNNs) have made considerable strides in several practical applications recently, including drug discovery, medical image processing, and recommendation. The graph structure of the GNN technique for recognizing diabetic retinopathy (DR) makes use of both labeled and unlabeled data. By using data relationships, managing unlabeled data, and scaling effectively, it improves accuracy. Despite their success, deep GNNs' performance is still constrained by several fundamental factors, one of which is over-smoothing. In deep Graph Neural Networks (GNNs), over-smoothing impairs generalization, reduces local structure capacity, impairs discrimination, and makes it difficult to capture higher-order connection patterns. Techniques like attentional processes and residual connections help to solve this problem. It shows that the stacked aggregators make the learned node representations almost impossible to discern. The performance deterioration of deep GNNs or their preference for feature correlation issues is examined. Performance problems with Deep Graph Neural Networks (GNNs) are brought on by feature correlation that results in noise, redundancy, and instability. To address these problems, preprocessing, feature selection, normalization, and regularisation approaches are required. To reduce the feature correlation issues, this work proposed DGCN for class-specific feature extraction from DR images. The Diabetic Retinopathy Grading Convolutional Network (DGCN), which is useful for automated diagnosis and grading, improves accuracy, lowers computational overhead, and enables robust feature extraction in diabetic retinopathy pictures. Figure 4 shows the DGCN feature extraction from SMOTE-balanced features.

3.2.1 GNN

The GNN approach moved the difficulty of DR recognition to semi-supervised learning on a graph by combining the properties of unlabelled vertices with those of neighboring labeled vertices. Due to its versatility, capacity for handling a variety of data, and capacity for capturing both local and global information, Graph Neural Networks (GNNs) are an excellent choice for diagnosing diabetic retinopathy. The GNN feature recognition model is presented in Figure 5, which is a graph convolution network. The eye fundus images are used in this instance as input in a connected graph method, where each node denotes the ideal feature, and each edge denotes the disease severity of the same feature.

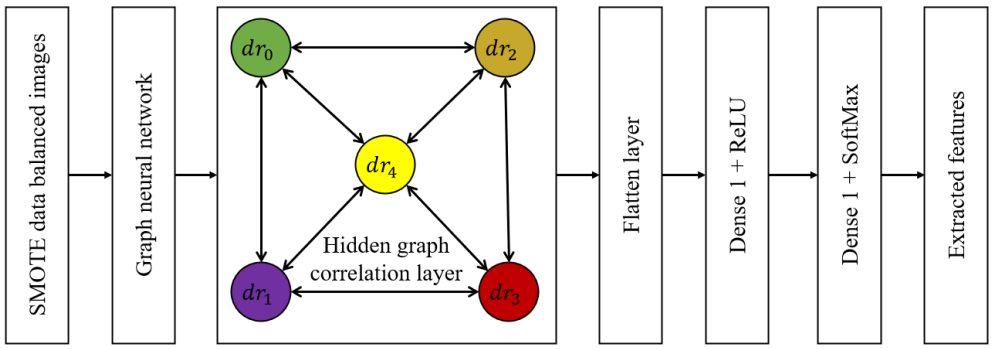


Figure 4: Proposed DGCN-based feature extraction.

The input for a connected graph is denoted by $G = (V, A)$, where V stands for nodes (vertex) and A for an adjacency matrix. Furthermore, the edge weights for all directions are stored in this adjacency matrix. The graph kernel and layer characteristics affect how well GNN performs. The individual layer operation of the GNN is shown in equation (3).

$$H^{l+1} = \sigma\left(D^{-0.5}AD^{-0.5}H^{(l)}W^{(l)}\right) \quad (3)$$

Here, l stands for the l^{th} layer, W for the linked graph's weight, H for its vertices, $\sigma(\cdot)$ represents the Leaky-ReLU activation function, and D for the diagonal matrix of $d_{i,j}$ components. When there is a lot of noise or outliers in the data: Leaky ReLU outperforms ReLU in situations when the data contains a lot of noise or outliers because it may produce a non-zero output for negative input values, preventing the loss of potentially crucial information. Additionally, the adjacency matrix element$a_{i,j}$ are used to create these $d_{i,j}$ elements:

$$d_{i,j}\text{s a} = \sum_{j=1}^{N} a_{i,j} \quad (4)$$

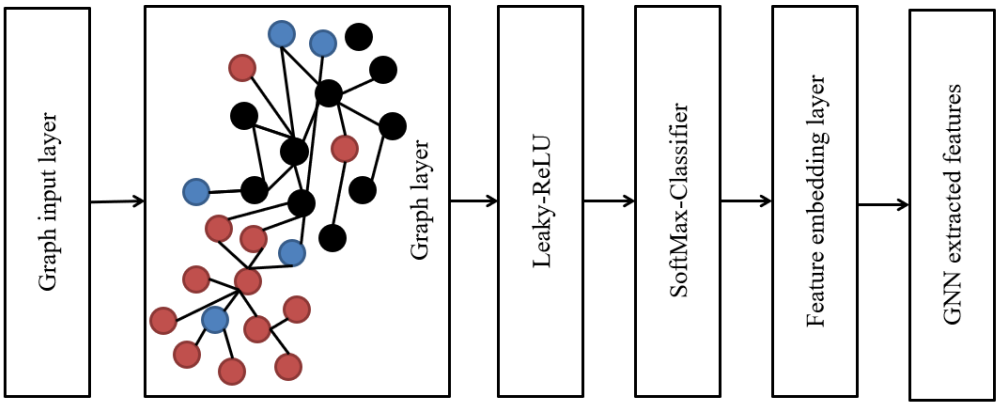


Figure 5: Architecture for GNN feature analysis.

The adjacency matrix and diagonal matrix must be constructed effectively to classify and grade DR using GNN. By recording spatial correlations and aiding feature propagation, the adjacency matrix is essential in Graph Neural Networks (GNNs) for diagnosing and grading diabetic retinopathy, increasing accuracy and interpretability. There are a total of five DR severity grades in the sample. As a result, each vertex in the linked graph layer bears a separate severity grade, and the adjacency matrix is designated as $A \in a_{5\times5}$. During training, an adjacency matrix is created based on the regularity of co-occurrences of various severity classes. In graph-based methods like Graph Neural Networks (GNNs), appearance probabilities are essential for directing connections and affecting graph sparsity. They boost GNN performance, information flow, and the graph's capacity to capture significant linkages. Additionally, each vertex is represented by a single hot encoding, allowing for simplex connections across graphs without any overlap. The adjacency matrix is produced during GNN training by computing the probabilistic co-occurrence probabilities. The appearance probability, $P_{ij} = P(L_i/L_j)$, is developed here by estimating the conditional probability ($P()$) between DR characteristics. Here, L_j stands for the likelihood of the test features, and L_i for the likelihood of the train features. By binarizing P_{ij}, the binary adjacency matrix (A') is created, improving the generalization stability of the model. Additionally, by taking into account the conditional probability-based weight bounds, A' is re-estimated, avoiding the over-smoothing issues brought on by features training. Here is how to represent the final conditional probability (A_f):

$$A'_{ij} = \begin{cases} 0, & P_{ij} < j \\ 1, & P_{ij} \geq j \end{cases} \quad (5)$$

$$A_f = \begin{cases} A'_{ij}, & i \neq j \\ 1 - P_{ij}, & i = j \end{cases} \quad (6)$$

Here, P_{ij} has a range of 0 to 1, depending on the random weight probability (τ). Additionally, at τ = 0.3, the optimal value of P is achieved as 0.25. Based on the updated A_f, the graph hidden layer operation is derived as follows:

$$H^{l+1} = \mathbb{f}\left(H^l, A_f\right) \quad (7)$$

$$\mathbb{f}\left(H^l, A_f\right) = \sigma\left(A_f, H^l, W^l\right) \quad (8)$$

In this case, the multi-grade feature separation is carried out using the feature matrix produced by the SoftMax classification operation, denoted by the symbol $\mathbb{f}(\cdot)$. A weight matrix for the lth neural network layer is called W^l.

3.2.2 Graph correlation layer

The GNN output features ($\mathbb{f}(H^l, A_f)$) were supplied to the graph correlation layer to identify the correlation among various features. Due to its effective use in medical data analysis, particularly for learning inherent correlations among various samples, DGCN is chosen to build learn graph feature representations and topological structure. To link each sample to the mini-batch of O_f is G_b, this research first creates graph correlations for the feature set $G_b = [G_1, \dots, G_k, \dots, G_B]$, where B was the batch size and $1 \leq k \leq B$. By creating an adjacent matrix A from this, DGCN links sample G_k with its K-nearest neighbors (KNN).

$$A_{kj} = \begin{Bmatrix} 0, if\ j \in K_k \\ 1, if\ j \in K_k \end{Bmatrix} \tag{9}$$

The K_k is a group of KNN of I_k, and A_{kj} stands for the correlation between the kth and jth samples. The graph correlations are developed by the mini batch $G(G_b, A)$ from the learned GNN feature G_b and neighboring matrix (A). One input layer, hidden graph correlation layers, and dense layers make up the DGCN used in this work. The hidden graph correlation layer operates in the manner described below.

$$X^{(l)} = \left[G_1^{(l)}, \dots, G_k^{(l)}, \dots, G_B^{(l)}\right].\sigma\left(D^{-1/2}AD^{-1/2}X^{(l-1)}W^{(l)}\right) \tag{10}$$

Here, the diagonal of the neighboring matrix is represented by D, and $X^{(l)}$ indicates the result of the hidden graph correlation layer with l samples.

$$D = diag(d_1, d_2, \dots, d_B) \tag{11}$$

$$d_B = \sum_{j=1}^{B} A_{kj} \tag{12}$$

The ReLU activation function is represented by $\sigma(.)$, the many diagonal components (d) that develop the D matrix. Then, under the restrictions of a graph-center loss, DGCN builds a KNN graph by calculating similarities to choose the top K samples. Similar features may be brought closer to one another, and discriminant correlations can be created by these methods for DGCN. This scenario shows the lack of correlation between head classes and tail classes during the dataset training". So, the hidden graph correlation layer effectively trains the features of head and tail classes by forming multi-level correlations. Finally, the DGCN consists of two dense layers with ReLU and SoftMax activation functions. These dense layers with activation functions formed as the fully connected layer. Here, the SoftMax classifier identifies the disease-specific features. The SoftMax classifier uses a labeled dataset to learn correlations between attributes and grades, giving test cases the grade with the best likelihood to help with multi-class grading problems like detecting diabetic retinopathy.

3.3 XGBoost classifier

As seen in Figure 6, the DGCN collected and balanced data samples are used as input to the XGBoost. In this case, to identify between the different assaults from test data, the XGBoost classifier is used. The training and testing process will employ the XGBoost ensembles of the XGBoost classifier. To deal with classification and regression problems, the XGBoost ensemble learning method, based on DT, is used. A decision tree ensemble learning method called XGBoost is applied to classification and regression problems. It is a potent machine-learning technique that predicts class labels and continuous target variables using gradient boosting and regularization. The well-known machine learning method XGBoost combines gradient boosting, parallelism, and sparsity awareness for resilience and flexibility. It effectively handles classification and regression problems. XGBoost falls under the umbrella of boosting algorithms since it strengthens a collection of poor learners over time. There are several advantages to using XGBoost. First, compared to previous boosting strategies, it learns from the data more quickly since it builds trees concurrently instead of consecutively.

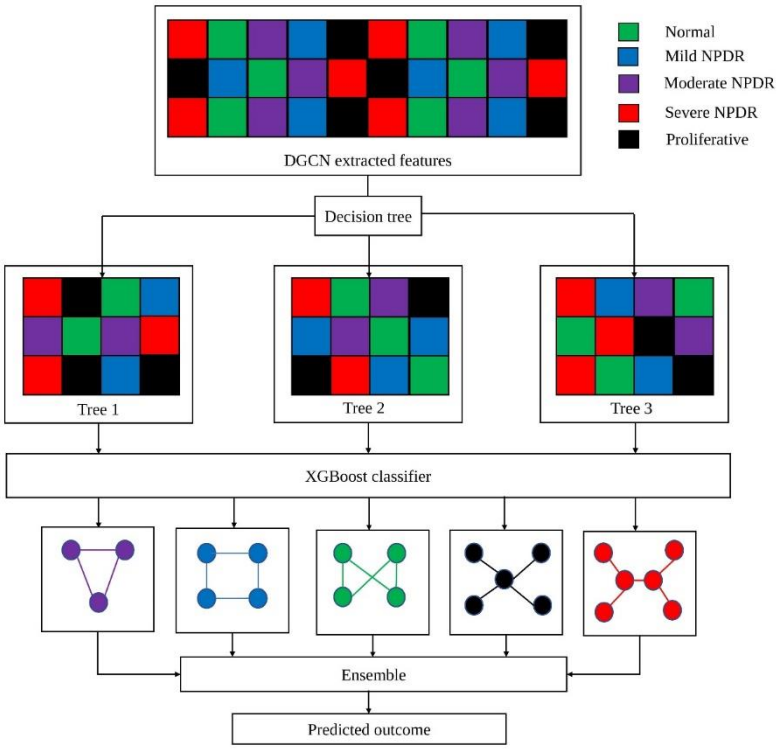


Figure 6: XGBoost classifier block diagram for obtaining prediction outcome.

Second, to reduce overfitting, XGBoost has a regularisation technique built in. Thirdly, it uses an approximation strategy to speed up the model training process. Therefore, XGBoost supports out-of-core processing and effectively manages weighted and sparse data. These elements have contributed to the XGBoost supervised learning algorithm, which is based on DT, becoming widely used. Due to its outstanding performance, efficiency, adaptability, resilience to overfitting, comprehension, community support, flexibility, and availability, XGBoost is a well-liked supervised learning strategy and a top pick among machine learning practitioners. Consider dataset, $D = \{(x_i, y_i)\}$, with $|D| = n, x_i \epsilon R^m$ and m features and n samples to illustrate how XGBoost works. The model output of a boosting technique utilizing T trees is defined as follows:

$$\hat{y}_i = \sum_{t=1}^{T} f_k(x_i),\ f_t \in \mathcal{F} \tag{13}$$

Where $F = \{f(x) = \omega_{II}(x)\}$ is a collection of trees constructed to address a classification problem. A tree is divided into two sections by each f_t: the leaf weight component (ω) and the structural part. With the use of objective functions like mean squared error or log-loss, the decision tree model's performance is assessed both before and after a node split. In XGBoost, the iterative procedure optimizes the tree structure to reduce loss and boost model precision. By reducing the following objective function, f_t may be discovered:

$$\mathcal{O} = \sum_i \ell(\hat{y}_i y_i) + \sum_t \Omega(f_t) \tag{14}$$

In this case, the distance between the prediction ($\hat{y}_i$) and object (y_i) is measured by a training loss function called ℓ, which stands for the cost of model complexity. Conventional methods fail to optimize a boosting algorithm with the aim function denoted by equation (4) in Euclidean space. In the gradient boosting approach, the prediction, and the objective function of the k-th iteration are specified as follows:

$$\hat{y}^{(k)} = \hat{y}^{(k-1)} + f_k(x) \tag{15}$$

$$\mathcal{O}^{(k)} = \sum_{i=1}^{n} \ell\left(y_i, \hat{y}_i^{(k-1)} + f_k(x_i)\right) + \Omega(f_k) \tag{16}$$

The second-order Taylor expansion is used by XGBoost to closely approach Equation (7). The goal function is expressed as follows:

$$\mathcal{O}^{(k)} \simeq \tilde{\mathcal{O}}^{(k)} = \sum_{i=1}^{n} \left[\ell\left(y_i, \hat{y}_i^{(k-1)}\right) + \}_i f_k(x_i) + \frac{||_i (f_k(x_i))^2}{2}\right] + \Omega(f_k) \tag{17}$$

Here k_i represents second-order gradient probability, which depends on the loss function, and $\mathcal{O}^{(k)}$ represents first-order gradient probabilities. Following is a definition of the XGBoost:

$$\Omega(f) = \gamma L + \frac{1}{2}\lambda \|\omega\|^2 \tag{18}$$

The number of leaves on the tree in this case is L. Take $\mathcal{L}_j = \{i: II(x_i) = j\}$ as an example of an instance asset. By adding a coefficient, equation (8) may now be readily expressed as Ω follows:

$$\tilde{\mathcal{O}}^{(k)} = \sum_{j=1}^{L}\left[\left(\sum_{i\in I_j}\}_i\right)\omega_j + \frac{1}{2}\left(\sum_{i\in I_j}||_i + \lambda\right)\omega_j^2\right] + \gamma L \quad (19)$$

The leaf j's solution weight (ω_j^*) for a tree structure may be calculated by:

$$\omega_j^* = -\frac{\sum_{i\in I_j}\}_i}{\sum_{i\in I_j}||_i+\lambda} \quad (20)$$

Equations (19) and (20) put together, we may create.

$$\tilde{\mathcal{O}}(II) = -\frac{1}{2}\sum_{j=1}^{L}\frac{\left(\sum_{i\in I_j}\}_j\right)^2}{\sum_{i\in I_j}||_j+\lambda} + \Upsilon L \quad (21)$$

Tree II(x) may be evaluated using equation (11) to find the best tree architectures. From a single leaf, the structure is gradually increased by adding branching after every repetition. A single leaf node is the starting point for decision tree algorithms like XGBoost, which divides it into two child nodes based on the cost or impurity measure. This procedure is performed several times, increasing the complexity of the tree, and limiting overfitting to produce an ensemble model. To add a specific split to the existing structure, the following method must be called:

$$\mathcal{O}_{split} = \frac{1}{2}\left[\frac{\left(\sum_{i\in I_l}\}_j\right)^2}{\sum_{i\in I_l}||_i+\lambda} + \frac{(\sum_{i\in I_r}\}_i)^2}{\sum_{i\in I_r}||_i+\lambda} + \frac{(\sum_{i\in I}\}_i)^2}{\sum_{i\in I}||_j+\lambda}\right] - \gamma \quad (22)$$

where $= I_l \cup I_r$, I_l, and I_r represent the instance sets of the left and right nodes after the split. The performance of the model as determined by the objective function is taken into account following a node split in the tree. If performance has improved, the appropriate split will be put into effect; if not, it will be stopped. Additionally, while enhancing the target function, XGBoost frequently encounters less overfitting than other boosting algorithms because of this regularisation. The target function, which specifies the desired prediction of the model, serves as the foundation for supervised learning. It directs the learning process by generating the necessary output for the inputs.

4 Results and discussion

This section gives a detailed analysis of simulation results performed using the proposed DRG-Net approach and existing DR grading classification methods which are verified using various parameters through the same dataset.

4.1 Dataset description

EyePACS dataset: There are 1427 images in the dataset. The training and validation sets were created by dividing the dataset's pictures 80:20. This division uses a random-split methodology throughout many training runs. A multi-stage training procedure was used to train several models. Multi-stage training includes simultaneously or sequentially training several models, assessing, and improving performance at each step, maybe modifying parameters, and applying ensemble approaches for improved accuracy. A training set was used for each run's training, and a validation set was used for each run's hyperparameter tuning.

Messidor dataset: The dataset contains 1744 fundus images with varied pixel sizes, which were obtained from three ophthalmologic departments. Due to characteristics including illness incidence, case diversity, clinical significance, patient consent, anonymization, standardization, and availability, the Messidor dataset is preferred for ophthalmological research. Stage 1 scans include those with MA. Stage 2 scans are those discovered in both MA and HM, while stage 3 scans are those discovered in high levels of MA and HM. Table 2 shows the dataset properties for the EyePACS and Messidor datasets. After applying the SMOTE data balancing, the dataset is normalized to equal numbers i.e., 25810 in EyePACS and 1017 in Messidor datasets.

Table 2: Number of instances in each dataset.

Dataset	**EyePACS**	**Messidor**
Class	No. of Instances	No. of Instances
Normal (No-DR)	25810	1017
Mild DR	2443	270
Moderate DR	5292	347
Severe DR	873	75
Proliferative DR	708	35

4.2 Prediction results

Figure 7 shows the predicted grading outcomes using the proposed DRG-Net. Here, the proposed method performs grading operations based on maximum feature-specific probability analysis. The suggested grading technique determines the grade with the highest likelihood based on each feature's specific contribution, allowing for

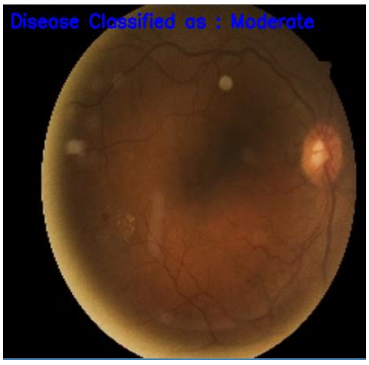
more precise and informed grading. These grades were identified by comparing the test features with the trained dataset using the SoftMax classifier. There are five distinct samples considered, and the yellow color-outlined values indicate the proposed DRG-Net probability estimate by the SoftMax classifier. Here, the DR grade is finalized based on the maximum value. For example, consider Figure 7 (a), which has probability values as **0.91,** 0.02, 0.03, 0.03,0.01. Here, **0.91** is the maximum achieved accuracy, which has the first position, so the output grade is classified as "No-DR".

0.91, 0.02, 0.03, 0.03,0.01

(a)

0.08**, 0.82,** 0.01, 0.06,0.03

(b)



0.04**,** 0.05**, 0.78**, 0.01,0.12

(c)

0.03, 0.06, 0.01, **0.74**,0.16

(e)

0.02, 0.08, 0.07, 0.03,**0.70**

(f)

Figure 7: Obtained results for DR grading prediction.

4.3 Ablation study

Ablation study is the process of performance estimation of DRG-Net in the presence and absence of various modules. The ablation research seeks to assess DRG-Net's constituent parts, such as feature importance, architecture choices, hyperparameters, loss functions, ensemble components, preprocessing techniques, data augmentation, regularisation techniques, and class balancing strategies. Table 3 and Table 4 show the ablation study of the proposed DRG-Net on both EyePACS and Messidor datasets. Here, the proposed method in the presence of all modules (SMOTE, DGCN, XGBoost) resulted in superior performance as compared to the absence of individual modules. Specifically, the absence of SMOTE data balancing and DGCN feature extraction resulted in poor performance. Combining real and fake data, changing the input layer, training, and assessing performance are the steps involved in incorporating balanced SMOTE data features into a dynamic graph convolutional system.

Table 3: Ablation study of proposed DRG-Net on EyePACS Dataset.

Presented modules	**Absence modules**	**Accuracy**	**Sensitivity**	**Specificity**
DGCN, XGBoost	SMOTE	93.01	92.35	96.12
SMOTE,	DGCN	97.54	93.48	97.45

XGBoost				
SMOTE, DGCN	XGBoost	97.89	94.78	98.12
SMOTE, DGCN, XGBoost	-	99.016	99.019	98.4375

Table 4: Ablation study of proposed DRG-Net on Messidor Dataset.

Presented modules	**Absence modules**	**Accuracy**	**Sensitivity**	**Specificity**
DGCN, XGBoost	SMOTE	90.35	93.19	95.61
SMOTE, XGBoost	DGCN	96.60	97.53	98.0
SMOTE, DGCN	XGBoost	97.34	96.35	96.45
SMOTE, DGCN, XGBoost	-	99.60	99.08	100.00

4.4 Performance evaluation

Table 5 shows the performance comparison of the proposed DRG-Net method with DCNN [12], CBAM [14], OCNN [16], and DBNN [20] using the EyePACS dataset. Here, the proposed method resulted in superior performance in terms of accuracy, sensitivity, and specificity, respectively. Table 6 shows the performance comparison of the proposed DRG-Net method with ERNN [18], Firefly-CNN [19], HA-Net [26], and MRCNN [30]. Here, the proposed method resulted in superior accuracy performance in contrast with other approaches using the Messidor dataset.

Table 5: Performance evaluation on EyePACS dataset.

Class	**Accuracy**	**Sensitivity**	**Specificity**
DCNN [12]	93.134	93.616	94.755
CBAM [14]	94.368	95.711	95.300
OCNN [16]	95.116	96.068	95.304
DBNN [20]	96.195	97.244	95.463
DRG-Net	99.016	99.019	98.4375

Table 6. Performance evaluation on Messidor dataset

Class	**Accuracy**	**Sensitivity**	**Specificity**
ERNN [18]	93.352	93.324	93.569
Firefly -CNN [19]	94.735	93.797	94.255
HA-Net [26]	96.511	95.389	94.322
MRCNN [30]	97.588	95.648	95.503
DRG-Net	99.60	99.08	100.00

5 Conclusion

This work developed DRG-Net by making use of the features of graph learning, and deep learning. Datasets, the datasets EyePACS and Messidor are subjected to the SMOTE algorithm, which brings the number of instances belonging to each DR class down to a standard level. So, the number of images in each class is maintained the same. After that, DGCN is used to extract the features that are unique to each class by locating the relationship between the classes. In addition, the prediction operation is carried out by an XGBoost classifier that has been trained with balanced feature sets. The results of the simulations performed on the EyePACS and Messidor datasets showed that the DRG-Net outperformed the more conventional methods in terms of performance. For improved performance, this work can be expanded with feature selection techniques.

Declarations

Funding

No funds or grants were received by any of the authors.

Conflict of interest

There is no conflict of interest among the authors.

Data Availability

All data generated or analysed during this study are included in the manuscript.

Code Availability

Not applicable.

Author's contributions

Venkata Kotam Raju Poranki and B. Srinivasa Rao. contributed to the design and methodology of this study, the assessment of the outcomes and the writing of the manuscript.

References

[1] Li, F., Wang, Y., Xu, T., Dong, L., Yan, L., Jiang, M., ... & Zou, H (2022). Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs. *Eye*, pp. 1433-1441. https://doi.org/10.1038/s41433-021-01552-8.

[2] AbdelMaksoud, E., Barakat, S., & Elmogy, M. (2022). A computer-aided diagnosis system for detecting various diabetic retinopathy grades based on a hybrid deep learning technique. *Medical & Biological Engineering & Computing*, pp. 2015-2038. https://doi.org/10.1007/s11517-022-02564-6.

[3] Vaghefi, E., Hill, S., Kersten, H. M., & Squirrell, D. (2020). Multimodal retinal image analysis via deep learning for the diagnosis of intermediate dry age-related macular degeneration: a feasibility study. *Journal of Ophthalmology*, Article ID 7493419. https://doi.org/10.1155/2020/7493419.

[4] Aziz, T., Ilesanmi, A. E., & Charoenlarpnopparut, C. (2021). Efficient and Accurate Hemorrhages Detection in Retinal Fundus Images Using Smart Window Features. *Applied Sciences*, pp. 6391. https://doi.org/10.3390/app11146391.

[5] Saranya, P., Prabakaran, S., Kumar, R., & Das, E (2022). Blood vessel segmentation in retinal fundus images for proliferative diabetic retinopathy screening using deep learning. *The Visual Computer*, pp. 977-992. https://doi.org/10.1007/s00371-021-02062-0.

[6] Zhao, L., Ren, H., Zhang, J., Cao, Y., Wang, Y., Meng, D., ... & Liu, F (2020). Diabetic retinopathy, classified using the lesion-aware deep learning system, predicts diabetic end-stage renal disease in Chinese patients. *Endocr Pract*, pp. 429–443. https://doi.org/10.4158/ep-2019-0512.

[7] Gao, Z., Li, J., Guo, J., Chen, Y., Yi, Z., & Zhong, J. (2019). Diagnosis of diabetic retinopathy using deep neural networks. *IEEE Access*, pp. 3360-3370. https://doi.org/10.1109/access.2018.2888639.

[8] Jabbar, M. K., Yan, J., Xu, H., Ur Rehman, Z., & Jabbar, A. (2022). Transfer learning-based model for diabetic retinopathy diagnosis using retinal images. *Brain Sci*, pp. 535. doi: 10.3390/brainsci12050535. PMID: 35624922; PMCID: PMC9139157. https://doi.org/10.3390/brainsci12050535.

[9] Farag, M. M., Fouad, M., & Abdel-Hamid, A. T (2022). Automatic severity classification of diabetic retinopathy based on DenseNet and convolutional block attention module. *IEEE Access*, pp. 38299-38308. doi: 10.1109/ACCESS.2022.3165193.

[10] Cahoon, S., Shaban, M., Switala, A., Mahmoud, A., & El-Baz, A. (2022). Diabetic retinopathy screening using a two-stage deep convolutional neural network trained on an extremely un-balanced dataset. *SoutheastCon2022*, pp. 250-254. doi: 10.1109/SoutheastCon48659.2022.9764079.

[11] Martinez-Murcia, F. J., Ortiz, A., Ramírez, J., Górriz, J. M., & Cruz, R. (2021). Deep residual transfer learning for automatic diagnosis and grading of diabetic retinopathy. *Neurocomputing (September 2021)*, pp. 452. 424-434. doi: 10.1016/j.neucom.2020.04.148.

[12] Majumder, S., & Kehtarnavaz, N. (2021). Multitasking deep learning model for detection of five stages of diabetic retinopathy. *IEEE Access*, pp. 123220-123230. doi: 10.1109/ACCESS.2021.3109240.

[13] Tavakoli, M., Mehdizadeh, A., Aghayan, A., Shahri, R. P., Ellis, T., & Dehmeshki, J. (2021). Automated microaneurysms detection in retinal images using radon transform and supervised learning: application to mass screening of diabetic retinopathy. *IEEE Access*, pp. 67302-67314. doi: 10.1109/ACCESS.2021.3074458.

[14] F Saeed, F., Hussain, M., & Aboalsamh, H. A. (2021). Automatic diabetic retinopathy diagnosis using adaptive fine-tuned convolutional neural network. *IEEE Access*, pp. 41344-41359. doi: 10.1109/ACCESS.2021.3065273.

[15] Raja Kumar, R., Pandian, R., Prem Jacob, T., Pravin, A., & Indumathi, P. (2021). Detection of diabetic retinopathy using deep convolutional neural networks. *Computational Vision and Bio-Inspired*

Computing, Springer, Singapore, pp. 415-430. https://doi.org/10.1007/978-981-33-6862-0_34.

[16] Gurani, V. K., Ranjan, A., & Chowdhary, C. L. (2019). Diabetic retinopathy detection using neural network. *International Journal of Innovative Technology and Exploring Engineering*, pp. 5. https://doi.org/10.35940/ijitee.j1105.0881019.

[17] Pao, S. I., Lin, H. Z., Chien, K. H., Tai, M. C., Chen, J. T., & Lin, G. M (2020). Detection of diabetic retinopathy using bichannel convolutional neural network. *Journal of Ophthalmology*, 2020. https://doi.org/10.1155/2020/9139713.

[18] Pires, R., Avila, S., Wainer, J., Valle, E., Abramoff, M. D., & Rocha, A (2019). A data-driven approach to referable diabetic retinopathy detection. *Artificial Intelligence in Medicine*, pp. 93-106. https://doi.org/10.1016/j.artmed.2019.03.009.

[19] Gadekallu, T. R., Khare, N., Bhattacharya, S., Singh, S., Maddikunta, P. K. R., Ra, I. H., & Alazab, M (2020). Early detection of diabetic retinopathy using PCA-firefly based deep learning model. *Electronics*, pp. 274. doi:10.3390/electronics9020274.

[20] Farooq, M. S., Arooj, A., Alroobaea, R., Baqasah, A. M., Jabarulla, M. Y., Singh, D., & Sardar, R (2022). Untangling computer-aided diagnostic system for screening diabetic retinopathy based on deep learning techniques. *Sensors, pp*. 1803. https://doi.org/10.3390/s22051803.

[21] Panwar, A., Semwal, G., Goel, S., & Gupta, S (2022). Stratification of the lesions in color fundus images of diabetic retinopathy patients using deep learning models and machine learning classifiers. *Edge Analytics*, Springer, Singapore, pp. 653-666. https://doi.org/10.1007/978-981-19-0019-8_49.

[22] Saini, M., & Susan, S (2022). Diabetic retinopathy screening using deep learning for multi-class imbalanced datasets. *Computers in Biology and Medicine*, pp. 105989. https://doi.org/10.1016/j.compbiomed.2022.105989.

[23] Kalpana Devi, M., & Mary Shanthi Rani, M (2022). Classification of diabetic retinopathy using ensemble of machine learning classifiers with IDRiD dataset. In *Evolutionary Computing and Mobile Sustainable Networks*, Springer, Singapore, pp. 291-303. https://doi.org/10.1007/978-981-16-9605-3_20.

[24] Shaukat, N., Amin, J., Sharif, M., Azam, F., Kadry, S., & Krishnamoorthy, S (2022). Three-Dimensional semantic segmentation of diabetic retinopathy lesions and grading using transfer learning. *Journal of Personalized Medicine*, pp. 1454. https://doi.org/10.3390/jpm12091454.

[25] Jang, S. I., Girard, M. J., & Thiery, A. H (2022). Explainable and interpretable diabetic retinopathy classification based on neural-symbolic learning. *arXiv preprint arXiv:2204.00624*, 2022.

[26] Shaik, N. S., & Cherukuri, T. K. (2022). Hinge attention network: A joint model for diabetic retinopathy severity grading. *Applied Intelligence*, pp. 15105-15121. https://doi.org/10.1007/s10489-021-03043-5.

[27] Da Rocha, D. A., Ferreira, F. M. F., & Peixoto, Z. M. A (2022). Diabetic retinopathy classification using VGG16 neural network. *Research on Biomedical Engineering*, pp. 761-772. https://doi.org/10.1007/s42600-022-00200-8.

[28] Yu, M., & Wang, Y. (2022). Intelligent detection and applied research on diabetic retinopathy based on the residual attention network. *International Journal of Imaging Systems and Technology*, pp. 1789-1800. https://doi.org/10.1002/ima.22734.

[29] Ganesh, M., Dulam, S., & Venkatasubbu, P. (2022). Diabetic Retinopathy Diagnosis with InceptionResNetV2, Xception, and EfficientNetB3. *In Artificial Intelligence and Technologies*, Springer, Singapore, pp. 405-413. https://doi.org/10.1007/978-981-16-6448-9_41.

[30] Erciyas, A., Barışçı, N., Ünver, H. M., & Polat, H. (2023). Improving detection and classification of diabetic retinopathy using CUDA and Mask RCNN. *Signal, Image, and Video Processing*, pp. 1265–1273. https://doi.org/10.1007/978-981-16-6448-9_41.

[31] Aiswarya, R. S., Kumar, V., & Punitha, P (2023). The Effect of Grain Size and Silicon Content on Non-Oriented Grain Steel Anomalous Loss Through Frequency Excitation in The Medical Healthcare by Using Big Data Analysis. *Tamjeed Journal of Healthcare Engineering and Science Technology*, pp. 43-53. https://doi.org/10.59785/tjhest.v1i1.5.